Supplementary Material text

The supplementary material text provides detailed description of the quantum and classical models, including derivations of the results in the main text and some additional details regarding the experimental materials.

Derivation leading to Equation (1)

This derivation explains the basic Quantum Zeno effect, under idealized conditions. The idealized situation referred to in the main text concerns a 2D quantum system, evolving under a unitary time independent Hamiltonian.

We prepare our system such that the initial state is $|I\rangle$ at t=0 and let it evolve for a total time $T > 0$. We are interested in the probability that measurements performed on the state at each of the times $T/_N$, $2T/_N$ ...T will confirm that the state is still $|I\rangle$. We have that:

$$Prob\left(I\ at\ time\,\frac{T}{N} \wedge\ I\ at\,\frac{2T}{N} \wedge \dots\right) = \left|\left(P_I e^{-\frac{iHT}{N}}\right)^N |I\rangle\right|^2 = \left|\left\langle I\left|e^{-\frac{iHT}{N}}\right|I\right\rangle\right|^{2N} \quad \text{(S1)}$$

For a two-level system and a time independent Hamiltonian, transition probabilities typically take the form $\left|\left\langle I\left|e^{-iHt}\right|I\right\rangle\right|^2 = cos^2(E \cdot t)$. In physical applications, E is usually an energy variable. Here, it can be thought of as the average strength of a piece of evidence, since Et is the rotation angle of the mental state, when presented with t pieces of evidence. Eq(S1) then readily leads to the expression, which is Eq(1) in the main text:

$$Prob\left(I\ at\ time\,\frac{T}{N} \wedge\ I\ at\,\frac{2T}{N} \wedge \dots\right) = cos^{2N}\left(\frac{\gamma}{N}\right),$$

where γ is a dimensionless constant.

Unitary dynamics and POVMs

In this section we motivate the particular choice of dynamics and measurement operators used in the quantum and Bayesian models. We will use this in the next section to derive Eq.(2), which is crucial in the present modeling, since it allows the setting of all parameters with classical data and thus prior to testing for the QZ effect.

In general, in situations such as the one we consider, the most appropriate form of dynamics would be non-unitary. This is because the expected evolution of the mental state is basically like a decay towards a fixed state, the guilty ray, since all the evidence participants encounter is that Smith is guilty and thus, asymptotically, participants must become certain that Smith is guilty.

However, there are two features of our experimental set up that mean that we never need consider mental states close to the guilty ray. First, all participants initially think Smith is innocent, and the evidence we present is designed to be weak, so that the probability that participants judge Smith to be guilty never rises above 50% (as evidenced in the data, e.g., see Figure 2). This means that the evolution by itself never leads to a state close to the guilty state. Thus, the only way a participant's mental state can end up close to the guilty state is by collapsing to this state, if the participant answers that Smith is guilty at one of the intermediate judgments. However, since our analyses were restricted to survival probability, we need not model the further evolution of the mental state after a guilty response. Thus, the only states whose dynamics we are interested in are those far from the guilty state. For these states the fact that the true evolution has a fixed point can, to a good approximation, be ignored, and so the dynamics of such states may be treated as unitary. Of course it is ultimately an empirical question whether this approximation allows for a good fit to the data. In addition, in future work, if it becomes relevant to explore a broader range of experimental manipulations within this paradigm and/or conditions for the mental state, then non-unitary dynamics could be employed.

So far, we have argued that we can model the dynamics of the cognitive state as unitary. However it turns out we need to consider time dependent unitary dynamics in order to capture the expected behavior of the cognitive state. This is essentially because we must allow for the fact that the 'strength' of a piece of evidence may depend on its serial position in the list of evidence presented. It is reasonable (especially in light of earlier remarks about the fact we expect the true evolution to have a fixed point) that we should expect to see a primacy effect, or equivalently diminishing returns, in the weight participants attach to different pieces of evidence. However when we explicitly introduce a form for the evolution in the next section we shall allow for the possibility of either a primacy or a recency effect, and leave it as an empirical question which behavior we see.

We also want to discuss the choice of POVMs to model the measurements. The particular POVMs we use simply model the impact of some noise on the measurements, so that the outcomes are no longer perfectly correlated with the cognitive state. Recall that the projectors representing Innocent and Guilty are given by $P_I = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, P_G = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$. The corresponding POVM operators that we use are $E_I = \begin{pmatrix} 1-\epsilon & 0 \\ 0 & \epsilon \end{pmatrix}, E_G = \begin{pmatrix} \epsilon & 0 \\ 0 & 1-\epsilon \end{pmatrix}$, where ϵ encodes the degree of noise. If a participant considers Smith innocent, so that the cognitive state is $|\psi\rangle = \binom{1}{0}$, then the probability of responding innocent is only $1-\epsilon$, leaving a probability to respond guilty of ϵ. Since ϵ is a parameter whose value we estimate from the data it may be that the best fit is provided by $\epsilon = 0$, in which case we recover the usual formalism of projective measurements. Note that the version of the collapse postulate that applies to POVMs is that after a measurement of the POVM E, which yields the answer 'yes', the state changes according to $|\psi\rangle \to \frac{\sqrt{E}|\psi\rangle}{|\sqrt{E}|\psi\rangle|}$. For more on POVMs see (26).

Derivation of Equation (2).

We can now proceed to derive Eq.(2) in the main text. At time 0 participants have not yet heard any evidence and at each time step participants are presented with evidence which supports the possibility of Smith's guilt. The probability that at $t = 0$ a participant initially responds that Smith is innocent is given as:

$$Prob(I\ at\ 0) = \langle\psi|E_I|\psi\rangle = (1-\epsilon)|\langle I|\psi\rangle|^2 + \epsilon|\langle G|\psi\rangle|^2 \qquad \text{(S2)}$$

where E_I is the POVM for innocent. This expression tells us that any participant who answers innocent for this initial judgment (before encountering any evidence) may be assumed to be in state $|I\rangle$ with probability $1-\epsilon$ and in state $|G\rangle$ with probability ϵ.

The general form of the transition probability for a time-dependent Hamiltonian is given by $Prob(I\ at\ time\ t) = \left|\left\langle I\left|e^{-i\int_0^t dsH(s)}\right|\psi\right\rangle\right|^2$. Then, the probability that a participant answers innocent after seeing t pieces of evidence, without any intermediate judgments, given an initial response of innocent, is

$$Prob(I\ at\ t|I\ at\ 0) = \frac{\left|\sqrt{E_I}e^{-i\int_0^t ds\, H(s)}\sqrt{E_I}|\psi\rangle\right|^2}{\left|\sqrt{E_I}|\psi\rangle\right|^2} \approx (1-\epsilon)\left|\sqrt{E_I}e^{-i\int_0^t ds\, H(s)}|I\rangle\right|^2$$

(S3)

To progress, we must make some assumptions regarding the Hamiltonian, $H(t)$. The Hamiltonian for any system in a two-dimensional Hilbert space can be written as a sum of the identity operator plus the three Pauli matrices, each with a time-dependent prefactor. As argued elsewhere (15, 18), it is reasonable to simplify the general expression for the time-dependent Hamiltonian of cognitive bivalued systems to $H(t) = b(t)\sigma_x = b(t)\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, where $b(t)$ is a function of time. Let us next define $B(t_m, t_n) = \int_{t_m}^{t_n} ds\, b(s)$, which incidentally is dimensionless. Then, Eq(S3) can be written as

$$
\begin{aligned}
Prob(I\ at\ t|I\ at\ 0) &= (1-\epsilon)\left|\sqrt{E_I}e^{-i\int_0^t ds\, H(s)}|1\rangle\right|^2 \\
&= (1-\epsilon)\left|\sqrt{E_I}\left(\mathrm{I}\cdot cos\big(B(0,t)\big) - i\sigma_x\, sin\big(B(0,t)\big)\right)|1\rangle\right|^2 \\
&= (1-\epsilon)^2 cos^2\big(B(0,t)\big) + \epsilon(1-\epsilon)sin^2\big(B(0,t)\big)
\end{aligned}
$$

which is Eq(2) in the main text.

Understanding the function $B(t_m, t_n)$, fixing it from data, and the Interpretation of the parameters

Both the quantum and classical models for opinion change involve the parameter ϵ, which takes into account erroneous responses, and the function $B(t_m, t_m)$, which tells us how the opinion state changes with accumulating evidence. In this section we describe how the function $B(t_m, n)$ can be specified, how to estimate it from empirical data, and how to interpret its parameters.

Recall, the function $B(t_m, t_n)$ controls the change of the mental state, as a result of considering $t_n - t_m$ pieces of evidence, assuming that a judgment was made at t_m. Therefore a naïve guess at this function would simply be the sum of the relative strengths of all pieces of evidence considered, multiplied by an overall constant, i.e.

$$B(t_m, t_n) =?\ \alpha \sum_{i=m+1}^{n} a_i$$

However the weight given to a piece of evidence may depend on its position in the sequence. Pieces of evidence that come later after a judgment may have less impact on the opinion state than pieces of evidence that come immediately after a judgment, or vice versa. Thus a better choice is,

$$B(t_m, t_n) = \alpha \sum_{i=m+1}^{n} a_i\, g(t_i - t_{m+1})$$

where the function $g(t)$ is a monotonic function of t. The choice of argument is made so that $B(0, t_1) = \alpha\, a_1 g(0)$, and we take $g(0) = 1$ by convention.

Note that the argument of $g(t)$ reflects the number of pieces of evidence seen since the last judgment was made, not the total number of pieces of evidence seen. This is very natural in the quantum model, since the idea is that the process of making a judgment 'collapses' the knowledge state back to the initial state (assuming an 'innocent' judgment.) This implies the state post-judgment should have the same sensitivity to evidence as the initial state, and so any primacy/recency effects should be reset. However this argument cannot be made in a Bayesian model, since 'collapse' is a characteristically quantum feature. Therefore the Bayesian model will involve a slightly different function, $B^C(t_m, t_n)$, where

$$B^C(t_m, t_n) = \alpha \sum_{i=m+1}^{n} a_i\, g(t_i - t_1)$$

There are many choices for the function $g(t)$. We will make the choice $g(t_i - t_{m+1}) = e^{-\beta(i-m-1)^2}$, so that overall we have:

$$B(t_m, t_n) = \sum_{i=m+1}^{n} \alpha\, a_i\ e^{-\beta(i-m-1)^2}$$

$$B^C(t_m, t_n) = \sum_{i=m+1}^{n} \alpha\, a_i\ e^{-\beta(i-1)^2} \qquad \text{(S4)}$$

A positive value of β corresponds to a primacy effect, or diminishing returns, whereas a negative value of β corresponds to a recency effect. This form for $g(t)$ may be motivated by considering a continuous analogue of the process of evidence presentation. Thus, our choice of $B(t_m, t_n)$, involves two free parameters, α, β. Note that there is no fitting regarding the relative strength parameters in Eq(S4), a_i. For a particular piece of evidence i, $a_i = \frac{average\ strength\ for\ evidence\ i}{average\ strength\ of\ all\ pieces\ of\ evidence}$, where both averages are across participants. Crucially the fact that we have reduced the determination of the functions $B(t_m, t_n)$ and $B^C(t_m, t_n)$ to the identification of two parameters means we can fix $B(t_m, t_n)$ and $B^C(t_m, t_n)$given data on $B(0, T)$, which in turn means we can fix it from data which does not concern intermediate judgments. The relative strength of the pieces of evidence, ie the a_i are given in Table 1S.

The parameter α is simply a factor that converts between evidence strength and angle of rotation of the opinion state. It is related to the overall strength of the prosecution's case, but it does not have a particularly interesting interpretation.

The parameter β is more interesting. Its inverse square root indicates the number of pieces of evidence after which the primacy or recency effect starts to have a large impact on the effect of additional evidence. For example, in Experiment 1, the best fit was for $\beta = 0.01$. This tells us that diminishing returns starts to play a role after around 10 pieces of evidence, so we would not expect to see much impact from this in the results. This is evident in Figure 2A, where we see a pure QZ effect. In contrast, in Experiment 2 the best fit was for $\beta = 0.0285$. This suggests diminishing returns should start to have an impact on behavior, after about 6 pieces of evidence. We can see this both in Figure 1B, where there is an obvious change in behavior from 6 to 12 pieces of evidence, and also in Figure 2B. In Figure 2B the noticeable dip in survival probability takes place between one judgment (i.e., only one judgment after all evidence has been presented) and two judgments. This is equivalent to considering the evidence either as one group of 12 pieces (evidence after 6 pieces would have a low impact, broadly speaking) or as two groups of 6 pieces of evidence (according to the quantum model, in this case, after 6 pieces of evidence and one judgment, the following 6 pieces of evidence would also be taken into account in the same way as the original 6; hence, the survival probability drops – more bias that Smith is guilty).

The best fit value for ϵ was approximately 3% in Experiment 1 and 1% in Experiment 2. This means that a participant whose cognitive state is perfectly aligned with the innocent ray may still have a $\approx$ 1% or 3% chance of answering that Smith is guilty, when queried. While this does not appear high for any individual judgment, in an experiment which employs more than two or three judgments, the cumulative error rate can quickly increase beyond 5%. Therefore, with multiple judgments, even in the presence of a simple procedure and very clear instructions (as in the present work), the possibility that participants respond incorrectly (i.e., in a way inconsistent with their mental state) needs to be incorporated in any modeling. The difference in the value of ϵ between Experiment 1 and Experiment 2 explains why there is a dip in survival probability for large N in Experiment 1 (Figure 2A) but this is not observed in Experiment 2 (Figure 2B).

Computing the (quantum) survival probability, for N intermediate measurements (Equation 3)

This section presents the derivation for the quantum survival probability. Following the usual convention in this work of denoting innocence with $|I\rangle$, we have that:

$$Prob('survival', N) = Prob\left(I\ at \frac{T}{N} AND\ I\ at \frac{2T}{N} AND \ldots I\ at\ T\right) \approx$$

$$(1-\epsilon)\left|\prod_{j=0}^{N-1}\sqrt{E_I}\exp\left(-iB\left(\frac{jT}{N},\frac{(j+1)T}{N}\right)\sigma_x\right)|I\rangle\right|^2 =$$

$$(1-\epsilon)\left|\prod_{j=0}^{N-1}\sqrt{E_I}\left(\mathrm{I}.\cos\left(B\left(\frac{jT}{N},\frac{(j+1)T}{N}\right)\right)-i\sigma_x.sin\left(B\left(\frac{jT}{N},\frac{(j+1)T}{N}\right)\right)\right)|I\rangle\right|^2$$

(S5)

These probabilities are quite complicated and it is not necessary to give the full expression for every value of N here. However, we can simplify them quite considerably by noting that both ϵ and $sin(B(t_i,t_j))$ are small compared to 1. Doing this allows us to write (this is Eq(3) in the main text):

$$\begin{aligned} Prob('survival', N) &= Prob\left(I\ at \frac{T}{N} AND\ I\ at \frac{2T}{N} AND \ldots I\ at\ T\right) \\ &= (1-\epsilon)^{N+1}\prod_{i=0}^{N-1}\cos^2\left(B\left(\frac{iT}{N},\frac{(i+1)T}{N}\right)\right) \\ &+ \epsilon(1-\epsilon)^N sin^2\left(B\left(\frac{(N-1)T}{N},T\right)\right)\prod_{i=0}^{N-2}\cos^2\left(B\left(\frac{iT}{N},\frac{(i+1)T}{N}\right)\right)+O(\epsilon^2) \\ &+ O(sin^4) \end{aligned}$$

(3)

Note that Eq(3) has a reasonably clear interpretation. The first term is the probability that the state never changes, multiplied by the probability that the N imperfect measurements all come out in the expected way (i.e., that Smith is innocent). The second term represents the probability that the state changes between the second to last and last measurements, but that the last measurement fails to detect this change. Further terms either represent earlier changes in the state, and so more failed detections, or the state changing back to innocent from guilty (the probability for this last possibility is expected to be negligible for other reasons, since a participant who thinks Smith is guilty is very unlikely to revert and respond that Smith is innocent, after seeing more guilty evidence).

Bayesian survival probability

To derive a Bayesian expression for survival probability, we will assume that the process of making a judgment does not affect the mental state, but, as judgments are imperfect, there is a small probability, ϵ, of making incorrect responses (that is, providing an answer which does not reflect the mental state).

As noted in the main text, much of the information we need to build a Bayesian model can be extracted from Eq(2). Recall that we denote by I_B the event where a participant *believes* Smith is innocent, and I_R the event where a participant *responds* that Smith is innocent, and similarly for guilty. Then from Eq(2) we have,

$$Prob(I_B\ at\ time\ t | I_B\ at\ time\ 0) = \cos^2(B(0,t))$$
$$Prob(G_B\ at\ time\ t | I_B\ at\ time\ 0) = \sin^2(B(0,t))$$
$$Prob(I_R|I_B) = (1-\epsilon), \qquad Prob(G_R\,|I_B\,) = \epsilon$$
$$Prob(G_R|G_B) = (1-\epsilon), \qquad Prob(I_R|G_B) = \epsilon$$

The probabilities involving transitions from Guilty cognitive states to Innocent ones are assumed to be 0, as in the quantum model.

The Bayesian survival probability is equal to,

$$Prob^C('survival', N) = prob\left(I_R \; at \; time \; T, I_R \; at \; time \frac{(N-1)T}{N}, \ldots I_R \; at \; time \frac{T}{N} \middle| I_R \; at \; 0\right)$$

We need two assumptions to allow us to write this in terms of quantities we know. The first is that ϵ is small, and the second is that transition probabilities from G_B to I_B are small. The first of these is justified by appeal to the data, the second by the nature of the empirical set up, since we only present evidence implying Smith's guilt. Given these two assumptions, we can show,

$$\begin{aligned} &Prob\left(I_R \; at \; time \; T, I_R \; at \; time \frac{(N-1)T}{N}, \ldots I_R \; at \; time \frac{T}{N} \middle| I_R \; at \; 0\right) \\ &\approx (1-\epsilon)^{N+1} \; Prob\left(I_B \; at \; time \; T, I_B \; at \; time \frac{(N-1)T}{N}, \ldots I_B \; at \; time \frac{T}{N} \middle| I_B \; at \; 0\right) \\ &+ \epsilon(1-\epsilon)^N \; Prob\left(G_B \; at \; time \; T, I_B \; at \; time \frac{(N-1)T}{N}, \ldots I_B \; at \; time \frac{T}{N} \middle| I_B \; at \; 0\right) \end{aligned}$$

This follows because the probability of transitioning back to I_B from G_B is essentially 0, and it is very unlikely that the state G_B is incorrectly classified by more than one judgment. Thus the only non-negligible possibility other than that the cognitive state was always aligned with innocent is that the state changed between the penultimate and final judgments.

Next, it is easy to see that,

$$\begin{aligned} &Prob(\ldots, I_B \; at \; time \; t_i, I_B \; at \; time \; t_{i-1}, \ldots I_B \; at \; time \; t_1 | I_B \; at \; 0) \\ &\approx Prob(\ldots, I_B \; at \; time \; t_i | I_B \; at \; 0), \end{aligned}$$

which follows because we are assuming the transition probabilities from G_B to I_B are small, so that if the state is I_B now, it is very unlikely to have been G_B at any time in the past. The survival probability then reduces to,

$$\begin{aligned} &Prob^C('survival', N) = \\ &\approx (1-\epsilon)^{N+1} \; Prob(I_B \; at \; time \; T | I_B \; at \; 0) \\ &+ \; \epsilon(1-\epsilon)^N \; Prob\left(G_B \; at \; time \; T, I_B \; at \; time \frac{(N-1)T}{N} \middle| I_B \; at \; 0\right) \end{aligned}$$

We can also write,

$$\begin{aligned} &Prob\left(G_B \; at \; time \; T, I_B \; at \; time \frac{(N-1)T}{N} \middle| I_B \; at \; 0\right) \\ &= Prob\left(G_B \; at \; time \; T \middle| I_B \; at \; \frac{(N-1)T}{N}\right) Prob\left(I_B \; at \; time \frac{(N-1)T}{N} \middle| I_B \; at \; 0\right) \end{aligned}$$

So we may finally write,

$$\begin{aligned} &Prob^C('survival', N) = \\ &\approx (1-\epsilon)^{N+1} \cos^2\left(B^C(0,T)\right) \\ &+ \; \epsilon(1-\epsilon)^N \sin^2\left(B^C\left(\frac{(N-1)T}{N}, T\right)\right) \cos^2\left(B^C\left(0, \frac{(N-1)T}{N}\right)\right) \end{aligned}$$

Additional details on the experimental methods.

Block	Evidence	Relative Strength, a_i	S.D.
1	Dixon was successful in his career and had recently been promoted.	0.92	0.49

	Dixon had arranged a number of social engagements for the week after his death.	0.83	0.48
	Dixon had no history of depression or related conditions.	0.94	0.48
2	Dixon was engaged to be married.	0.89	0.49
	One of Smith's previous housemates reported that Smith made him feel threatened.	1.15	0.50
	Friends and colleagues reported that Dixon did not seem obviously stressed or depressed in the days leading up to his death.	0.90	0.48
3	Neighbours reported overhearing Dixon and Smith engaged in heated conversations on the evening before Dixon's death.	1.25	0.43
	Dixon appeared to have a large quantity of savings.	0.70	0.46
	Smith had a previous conviction for assault.	1.22	0.44
4	Smith's fingerprints were found on the bottle of liquor, although it was impossible to tell whether these were recent.	1.01	0.53
	The addition of the sleeping pills to the liquor was unlikely to have altered its taste.	0.92	0.51
	The local pharmacist testified that Smith had bought the sleeping pills in his pharmacy recently after complaining of insomnia.	1.29	0.48

Table S1. The 12 pieces of evidence suggesting that Smith is guilty, with average relative strengths and standard deviations. This data was based on participants' judgments about the strength of evidence, as collected at the end of Experiments 1, 2. The average relative strength of evidence in blocks 1,2,3 and 4 is 0.90, 0.98, 1.06 and 1.07 respectively.

Details of the Bayesian Analyses

The computations of BIC and Bayes Factors were carried out following Jarosz and Wiley (22). In particular, the BIC was estimated from the R^2 via,

$$BIC = n * \ln(1 - R^2) + k * \ln(n)$$

Where k is the number of free parameters and n is the sample size. The Bayes factors were then computed in the usual way,

$$BF_{QB} = e^{\Delta BIC_{QB}/2}$$

where $\Delta BIC_{QB} = BIC_Q - BIC_B$ is the difference in BIC values for the Quantum and Bayesian models.

Additional references for Supplementary Materials

(26) Yearsley, JM and Busemeyer, JR (in press). Quantum cognition and decision theories: A tutorial. *Journal of Mathematical Psychology.*